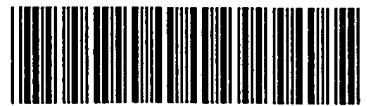# RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: _/O/566,886_

Source: _TFWP_

Date Processed by STIC: _2-/0-06_

ENTERED

RAW SEQUENCE LISTING   DATE: 02/10/2006
PATENT APPLICATION: **US/10/566,886**   TIME: 09:20:25

Input Set : **A:\14028.0295U2.txt**
Output Set: **N:\CRF4\02102006\J566886.raw**

```
 4 <110> APPLICANT: NEVILLE, David
 5       WOO, Jung-Hee
 6       LIU, Yuan-Yi
 8 <120> TITLE OF INVENTION: METHODS FOR EXPRESSION AND PURIFICATION
 9       OF IMMUNOTOXINS
11 <130> FILE REFERENCE: 14028.0295U2
13 <140> CURRENT APPLICATION NUMBER: US/10/566,886
14 <141> CURRENT FILING DATE: 2006-02-01
16 <150> PRIOR APPLICATION NUMBER: PCT/US04/24786
17 <151> PRIOR FILING DATE: 2004-08-02
19 <150> PRIOR APPLICATION NUMBER: 60/491,923
20 <151> PRIOR FILING DATE: 2003-08-01
22 <160> NUMBER OF SEQ ID NOS: 35
24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 22
28 <212> TYPE: PRT
29 <213> ORGANISM: H. sapiens
31 <400> SEQUENCE: 1
32 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
33 1               5                   10                  15
34 Ile Pro Thr Ala Arg Arg
35             20
37 <210> SEQ ID NO: 2
38 <211> LENGTH: 22
39 <212> TYPE: PRT
40 <213> ORGANISM: M. musculus
42 <400> SEQUENCE: 2
43 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
44 1               5                   10                  15
45 Ile Pro Thr Ala Arg Arg
46             20
48 <210> SEQ ID NO: 3
49 <211> LENGTH: 22
50 <212> TYPE: PRT
51 <213> ORGANISM: R. norvegicus
53 <400> SEQUENCE: 3
54 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
55 1               5                   10                  15
56 Ile Pro Thr Ala Arg Arg
57             20
59 <210> SEQ ID NO: 4
60 <211> LENGTH: 22
```

C--> (lines 13, 14)

RAW SEQUENCE LISTING   DATE: 02/10/2006
PATENT APPLICATION: US/10/566,886   TIME: 09:20:25

        Input Set : A:\14028.0295U2.txt
        Output Set: N:\CRF4\02102006\J566886.raw

61 <212> TYPE: PRT
63 <213> ORGANISM: C. griseus
65 <400> SEQUENCE: 4
66 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
67 1   5   10   15
68 Ile Pro Thr Ala Arg Arg
69   20
71 <210> SEQ ID NO: 5
72 <211> LENGTH: 22
73 <212> TYPE: PRT
74 <213> ORGANISM: D. melanogaster
76 <400> SEQUENCE: 5
77 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
78 1   5   10   15
79 Ile Pro Thr Thr Arg Arg
80   20
82 <210> SEQ ID NO: 6
83 <211> LENGTH: 22
84 <212> TYPE: PRT
85 <213> ORGANISM: C. elegans
87 <400> SEQUENCE: 6
88 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
89 1   5   10   15
90 Ile Pro Thr Ala Arg Arg
91   20
93 <210> SEQ ID NO: 7
94 <211> LENGTH: 22
95 <212> TYPE: PRT
96 <213> ORGANISM: S. pombe
98 <400> SEQUENCE: 7
99 Asp Val Val Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile
100 1   5   10   15
101 Ile Pro Thr Ala Arg Arg
102   20
104 <210> SEQ ID NO: 8
105 <211> LENGTH: 22
106 <212> TYPE: PRT
107 <213> ORGANISM: P. pastoris
109 <400> SEQUENCE: 8
110 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Val
111 1   5   10   15
112 Ile Pro Thr Met Lys Arg
113   20
115 <210> SEQ ID NO: 9
116 <211> LENGTH: 22
117 <212> TYPE: PRT
118 <213> ORGANISM: S. cerevisiae
120 <400> SEQUENCE: 9
121 Asp Val Thr Leu His Ala Asp Ala Ile His Arg Gly Gly Gly Gln Ile

RAW SEQUENCE LISTING   DATE: 02/10/2006
PATENT APPLICATION: US/10/566,886   TIME: 09:20:25

        Input Set : A:\14028.0295U2.txt
        Output Set: N:\CRF4\02102006\J566886.raw

```
122  1              5                10                    15
123 Ile Pro Thr Met Arg Arg
124              20
126 <210> SEQ ID NO: 10
127 <211> LENGTH: 66
128 <212> TYPE: DNA
129 <213> ORGANISM: Artificial Sequence
131 <220> FEATURE:
132 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
133       synthetic construct
135 <400> SEQUENCE: 10
136 gatgttaccc tgcacgccga tgctatccac cgccgcggag gacaagtcat tccaaccatg      60
137 aagaga                                                                 66
139 <210> SEQ ID NO: 11
140 <211> LENGTH: 223
141 <212> TYPE: DNA
142 <213> ORGANISM: Artificial Sequence
144 <220> FEATURE:
145 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
146       synthetic construct
148 <400> SEQUENCE: 11
149 actttgaagt tcttaatttt gttcctcgta gaaagaacgc atagataatt caaaatggca      60
150 aaatgggtat gtgttttttt atagttcatg tgccgaacaa ctaccgtttt aacttcactg     120
151 tcgatcagat gcgatccctt atggacaagg tgtccaacgt ccgtaacatg tcggttattg     180
152 cccacgttga tcacggtaag tccactttaa ctgactccct ggt                       223
154 <210> SEQ ID NO: 12
155 <211> LENGTH: 250
156 <212> TYPE: DNA
157 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
161       synthetic construct
163 <400> SEQUENCE: 12
164 actttgaagt tcttaatttt gttcctcgta gaaagaacgc atagataatt caaaatgggt      60
165 atgtgttttt ttatagttca tgtgccgaac aactaccgtt tcaagatggg agccagccac     120
166 taacatctcc tctagttaac ttcactgtcg atcagatgcg atcccttatg gacaaggtga     180
167 ccaacgtccg taacatgtcg gttattgccc acgttgatca cggtaagtcc actttaactg     240
168 actccctggt                                                             250
170 <210> SEQ ID NO: 13
171 <211> LENGTH: 2601
172 <212> TYPE: DNA
173 <213> ORGANISM: Artificial Sequence
175 <220> FEATURE:
176 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
177       synthetic construct
179 <400> SEQUENCE: 13
180 atggttaact tcactgtcga tcagatgcga tcccttatgg acaaggtgac caacgtccgt      60
181 aacatgtcgg ttattgccca cgttgatcac ggtaagtcca ctttaactga ctccctggtg     120
182 caacgtgccg gtattatttc tgctgccaag gctggtgagg cccgtttcac tgatactaga     180
```

```
183 aaggacgagc aagagagagg tatcaccatc aagtctaccg ccatttcttt gtactctgag   240
184 atgggtgacg acgatgtcaa ggagatcaag cagaagactg aaggtaacag tttccttatc   300
185 aacttaattg actccccagg tcacgttgac ttctcttctg aggtcactgc tgctctgcgt   360
186 gttactgacg gtgctttggt cgtcgttgac tgtgttgaag gtgtctgtgt tcaaactgag   420
187 accgttttgc gtcaagcttt gggtgaaaga atcaagccag ttgttgtcat taacaaggtc   480
188 gaccgtgctc ttttggagtt gcaagttacc aaggaggacc tgtaccagtc tttcgctaga   540
189 accgtcgagt ccgtaaacgt cgttatcgct acttacactg acaagaccat tggtgacaac   600
190 caagtctacc cagaacaggg taccgtcgct ttcggttcag gtctgcacgg atgggctttc   660
191 accgttagac agttcgccac tagatactcc aagaagttcg gtgttgacag aatcaagatg   720
192 atggagcgtc tgtggggaga ctcttacttc aacccaaaga ccaagaaatg gaccaacaag   780
193 gacaaggacg ccgctggaaa gcctttggag cgtgccttca acatgttcgt tttggaccct   840
194 atcttccgtc tgtttgctgc catcatgaac ttcaagaagg atgaaattcc agttctgttg   900
195 gagaaattgg agatcaacct gaagcgtgag gagaaggagt tggagggtaa ggctcttttg   960
196 aaggttgtca tgagaaagtt cttgccagct gccgacgctt tgttggagat gattgttctt  1020
197 cacctgccat ctccagtcac cgctcaagct tacagagccg agactttgta cgaaggtcca  1080
198 tctgatgacc aattctgcat tggtatcaga gagtgtgacc ctaaggctga gctgatggtt  1140
199 tacatttcca agatggtgcc aacctccgac aaaggtagat tctacgcctt cggtcgtgtt  1200
200 ttctccggta ctgttaagtc cggtcaaaag gtcagaatcc aaggtcctaa ctacgttcca  1260
201 ggtaagaagg aggacttgtt catcaaggct gttcaaagaa ctgttttgat gatgggaaga  1320
202 accgtcgagc ctattgacga tgtcccagct ggtaacattc tgggtattgt gggtatcgac  1380
203 cagttcttgc tgaagtctgg tactcttact accaacgaag ccgctcacaa catgaaggtg  1440
204 atgaaattct ctgtctctcc agttgtgcaa gttgccgttg aggtcaagaa cgctaatgat  1500
205 ctgcccaagt tggttgaggg tctgaagcgt ttgtccaagt ctgacccatg tgttttaacc  1560
206 tacatctccg agtctggtga gcacattgtt gctggtactg gtgagctgca cttggaaatc  1620
207 tgtttgcaag atctgcaaga cgaccacgct ggtgtccctc tgaagatttc tcctccagtt  1680
208 gttacctacc gtgagactgt cactaacgaa tcttccatga ctgccctgtc caagtctcag  1740
209 aacaagcata acagaattta cctgaaggct caaccaattg acgaggaatt gtctttggct  1800
210 atcgaagaag gtaaggttca cccaagagac gactttaaag ccagagccag aatcatggct  1860
211 gatgaatacg gttgggacgt cactgatgcc agaaagatct ggtgtttcgg tccagacggt  1920
212 actggtgcca acttagttgt tgaccagtct aaggctgtcc aatacttgca cgagatcaag  1980
213 gactctgttg ttgccggttt ccaattggct accaaggaag gtccaatttt gggagaaaac  2040
214 atgagatccg tcagagtcaa catcttggat gttaccctgc acgccgatgc tatccacaga  2100
215 ggtggaggac aagtcattcc aaccatgaag agagttacct acgccgcctt cctgttggct  2160
216 gagccagcta tccaggagcc tatcttcttg gtggagatcc aatgtccaga gaatgccatt  2220
217 ggtggtatct actctgtttt gaacaagaag agaggtcaag ttatctctga ggaacaaaga  2280
218 ccaggtaccc cattgttcac tgtcaaagct tacttgccag ttaacgagtc attcggtttc  2340
219 accggtgaac tgagacaagc taccgctggt caagctttcc cacagatggt gttcgaccac  2400
220 tgggccaaca tgaatggtaa cccattggac ccagcctcca aggtcggtga gattgttctt  2460
221 gctgccagaa agagacaggg tatgaaggag aacgttcctg gttatgaaga gtactacgac  2520
222 aagttgtaag cttaatgttt cattaactta tttgtgtcgt tcgtatgtct atttacgtac  2580
223 ttaattcagt gtattgttgt t                                            2601
225 <210> SEQ ID NO: 14
226 <211> LENGTH: 9
227 <212> TYPE: PRT
228 <213> ORGANISM: Artificial Sequence
230 <220> FEATURE:
231 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
232       synthetic construct
234 <400> SEQUENCE: 14
```

Input Set : **A:\14028.0295U2.txt**
Output Set: **N:\CRF4\02102006\J566886.raw**

```
235 Ala His Val Asp His Gly Lys Ser Thr
236 1               5
238 <210> SEQ ID NO: 15
239 <211> LENGTH: 13
240 <212> TYPE: PRT
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
245       synthetic construct
247 <400> SEQUENCE: 15
248 Asp Glu Gln Glu Arg Gly Ile Thr Ile Lys Ser Thr Ala
249 1               5                   10
251 <210> SEQ ID NO: 16
252 <211> LENGTH: 896
253 <212> TYPE: PRT
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: Description of Artificial Sequence; note =
258       synthetic construct
260 <400> SEQUENCE: 16
261 Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu
262 1               5                   10                  15
263 Asn Phe Ala Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile
264             20                  25                  30
265 Gln Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr Asp
266         35                  40                  45
267 Asp Asp Trp Lys Gly Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala Ala
268     50                  55                  60
269 Gly Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly Gly
270 65                  70                  75                  80
271 Val Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu Lys
272                 85                  90                  95
273 Val Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu Thr
274             100                 105                 110
275 Glu Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg Phe
276         115                 120                 125
277 Gly Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu Gly
278     130                 135                 140
279 Ser Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala Leu
280 145                 150                 155                 160
281 Ser Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly Gln
282                 165                 170                 175
283 Asp Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg Val
284             180                 185                 190
287 Arg Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp
288         195                 200                 205
289 Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His
290     210                 215                 220
291 Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Ala Lys Thr Val Ser
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 02/10/2006
PATENT APPLICATION: **US/10/566,886**   TIME: 09:20:26

Input Set : **A:\14028.0295U2.txt**
Output Set: **N:\CRF4\02102006\J566886.raw**

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:19; Xaa Pos. 2,3
Seq#:35; N Pos. 7,15

VERIFICATION SUMMARY DATE: 02/10/2006
PATENT APPLICATION: **US/10/566,886** TIME: 09:20:26

Input Set : **A:\14028.0295U2.txt**
Output Set: **N:\CRF4\02102006\J566886.raw**

```
L:13 M:270 C: Current Application Number differs, Replaced Current Application Number
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:432 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:626 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:35
L:631 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:0
```